3A91-23

A NOTE ON CONSTRUCTING SIMULTANEOUS CONFIDENCE INTERVALS ABOUT MULTINOMIAL PROBABILITIES

By

Jeffrey F. Bromaghin

Regional Information Report[1] NO. 3A91-23

Alaska Department of Fish and Game
Division of Commercial Fisheries, AYK Region
333 Raspberry Road
Anchorage, Alaska 99518

December, 1991

[1]The Regional Information Report Series was established in 1987 to provide an information access system for all unpublished divisional reports. These reports frequently serve diverse ad hoc informational purposes or archive basic uninterpretted data. To accommodate timely reporting of recently collected information, reports in this series undergo only limited internal review and may contain preliminary data; this information may be subsequently finalized and published in the formal literature. Consequently, these reports should not be cited without prior approval of the author or the Division of Commercial Fisheries.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The multinomial distribution is frequently used to model naturally occurring populations. For example, fisheries biologists often assume a multinomial model is appropriate when estimating the sex and age composition of fish populations. In many cases, sample data are used to construct interval estimates of the multinomial probabilities. Although the properties of individual confidence intervals may be of interest, it is equally common for researchers to be concerned with the properties of the collection of individual confidence intervals. In such cases, confidence intervals are constructed about the individual probabilities in such a way that the confidence in the collection of intervals is equal to or greater than some specified level. Such confidence intervals are termed simultaneous confidence intervals (Johnson and Bhattacharyya, 1987).

Goodman (1965) demonstrated that simultaneous confidence limits can be obtained as the solutions of a set of quadratic equations in the multinomial probabilities. Two methods of constructing simultaneous confidence intervals about multinomial probabilities, i.e., two sets of equations, were presented. Although one of the methods is generally acknowledged as the preferred method, e.g., Seber (1982), it appears to be used infrequently in practice. This work examines the properties of the two methods using computer simulation. The simulation results indicate that the less frequently used method consistently produces superior simultaneous confidence intervals.

NOTATION

Before discussing the two methods, some notation is required. Since the methods to be discussed are due to Goodman (1965), his notation is followed closely. Let

k	=	the number of categories in the multinomial distribution,
n	=	the sample size,
n_i	=	the number of the sampled units classified as belonging to the i^{th} category,
π_i	=	the probability a randomly selected unit would be classified as belonging in the i^{th} category,
$\hat{\pi}_i$	=	the maximum likelihood estimator of π_i, i.e., n_i/n,
π_i^-	=	the lower confidence limit for the i^{th} probability, and
π_i^+	=	the upper confidence limit for the i^{th} probability.

SIMULTANEOUS CONFIDENCE INTERVALS

Goodman (1965) demonstrated that $(1 - \alpha)100\%$ simultaneous confidence limits can be obtained as the roots of

$$(\hat{\pi}_i - \pi_i)^2 = z^2_{(1-\frac{\alpha_i}{2})} V , \tag{1}$$

$i = 1, 2, \ldots, k$, where the α_i sum to α and the quantity V is to be specified. Two choices of V are immediately obvious and were presented by Goodman (1965). Letting V be the variance of the random variable $\hat{\pi}_i$, we have

$$(\hat{\pi}_i - \pi_i)^2 = z^2_{(1-\frac{\alpha_i}{2})} \left(\frac{\pi_i(1-\pi_i)}{n} \right) . \tag{2}$$

It is easy to show that the roots of equation (2) are

$$\begin{aligned} \pi_i^+ &= \frac{z^2_{(1-\frac{\alpha_i}{2})} + 2n_i + z_{(1-\frac{\alpha_i}{2})} \sqrt{z^2_{(1-\frac{\alpha_i}{2})} + 4n_i \left(\frac{n - n_i}{n} \right)}}{2\left(n + z^2_{(1-\frac{\alpha_i}{2})} \right)} \\ \pi_i^- &= \frac{z^2_{(1-\frac{\alpha_i}{2})} + 2n_i - z_{(1-\frac{\alpha_i}{2})} \sqrt{z^2_{(1-\frac{\alpha_i}{2})} + 4n_i \left(\frac{n - n_i}{n} \right)}}{2\left(n + z^2_{(1-\frac{\alpha_i}{2})} \right)} \end{aligned} \tag{3}$$

(Goodman, 1965). The construction of simultaneous confidence intervals using the confidence limits of equation (3) will be referred to as Method 1.

Letting V be the maximum likelihood estimator of the variance of $\hat{\pi}_i$, we have

$$(\hat{\pi}_i - \pi_i)^2 = z^2_{(1-\frac{\alpha_i}{2})} \left(\frac{\hat{\pi}_i(1-\hat{\pi}_i)}{n-1} \right) . \tag{4}$$

The roots of equation (4) are the commonly used simultaneous confidence limits

$$\begin{aligned} \pi_i^+ &= \hat{\pi}_i + z_{(1-\frac{\alpha_i}{2})} \sqrt{\frac{\hat{\pi}_i(1-\hat{\pi}_i)}{n-1}} \\ \pi_i^- &= \hat{\pi}_i - z_{(1-\frac{\alpha_i}{2})} \sqrt{\frac{\hat{\pi}_i(1-\hat{\pi}_i)}{n-1}} \end{aligned} \tag{5}$$

(Goodman, 1965). The construction of simultaneous confidence intervals using the confidence limits of equation (5) will be referred to as Method 2.

If $k = 2$, a slight modification to the above procedure is warranted. In this case, confidence limits for one of the probabilities, say π_1, are constructed using $\alpha_1 = \alpha$ and either of the above methods. However, confidence limits for π_2 are constructed as

$$\begin{aligned} \pi_2^- &= 1 - \pi_1^+ \\ \pi_2^+ &= 1 - \pi_1^- . \end{aligned} \tag{6}$$

These confidence limits for π_2 have the same properties as do the limits for π_1 and the desired simultaneous confidence level is achieved. In addition, these confidence intervals are shorter than the limits computed using either of the above methods. This procedure is not advisable for $k \geq 3$ (Goodman, 1965).

THE SIMULATION STUDY

A simulation study was conducted to numerically compare the properties of the two methods of constructing simultaneous confidence intervals. The simulation was performed on a personal computer using a program written in C (Kernighan and Ritchie, 1988) and compiled with the Turbo C, version 2.0, compiler (Borland International, 1988). Uniform random deviates, used in sampling from multinomial distributions, were generated using the "ranl" function of Press, et al. (1986). Normal percentiles were computed using the "GAUINV" algorithm (Kennedy and Gentle, 1980).

Samples of size 25, 50, 75, 100, 150, 200, and 300 were drawn from a total of 8 multinomial distributions. The probabilities of the multinomial distributions sampled are given in Table 1. Five thousand independent samples of a given size were drawn from each multinomial distribution. For each sample, 95% simultaneous confidence intervals were constructed about the multinomial probabilities using both Method 1 and Method 2. In all cases, α_i took the value α/k for all i; employing the Bonferroni inequality (Sokal and Rohlf, 1987).

For the purposes of summarizing the results of these simulations, "individual coverage" is defined as the proportion of the samples in which the confidence interval captured the probability being estimated and "simultaneous coverage" is defined as the proportion of the samples in which the collection of confidence intervals captured all of the probabilities of the distribution. Based on the results observed in the 5000 samples, three criterion were used in comparing the characteristics of the methods; (1) the simultaneous coverage, (2) the individual coverage, and (3) the observed mean width of confidence intervals.

SIMULATION RESULTS

The simultaneous coverage values observed in each of the simulations are presented in Table 2. An examination of these results suggests that when the probabilities of the multinomial distribution are unequal, the simultaneous confidence intervals produced by Method 1 approach the desired simultaneous coverage more quickly, as sample size increases, than do the confidence intervals of Method 2. In fact, it appears that Method 2 often only approaches the desired confidence level asymptotically. A typical example of this observation is presented in Figure 1. When the probabilities of the multinomial distribution are equal, this pattern is not always observed (Figure 2). However, as the number of categories in the distribution increases, i.e., the probabilities decrease in magnitude, the pattern becomes increasingly apparent (Figure 3).

The observed individual coverage values and mean confidence interval widths are presented in Table 3. An examination of these results immediately reveals the reason that the above pattern is observed. For small probabilities, the confidence intervals produced by Method 2 do not achieve the desired individual coverage unless large samples are drawn (Figure 4). The confidence intervals produced by Method 1 achieve the desired individual coverage at smaller sample sizes (Figure 5). In other words, the confidence intervals of Method 2 are too narrow and frequently fail to capture small probabilities, causing the simultaneous coverage to fall below the specified minimum.

It is interesting to note that the confidence intervals produced by Method 1 are nearly always more narrow than the intervals produced by Method 2 (Table 3). In fact, the only time the confidence intervals of Method 2 are more narrow than the intervals produced by Method 1 is when the intervals of Method 2 are too narrow to capture the probabilities with sufficient frequency (Figure 6). Whenever both methods achieve the desired individual confidence level, Method 1 produces shorter confidence intervals.

DISCUSSION

The results of the simulation study indicate that Method 1 consistently produces simultaneous confidence intervals that are superior to those produced by Method 2. When the probabilities of the multinomial distribution are relatively large, both methods produce satisfactory parameter coverage; however, in this case, the confidence intervals of Method 2 are wider than those of Method 1. In addition, for small sample sizes, the confidence intervals produced by Method 2 fail to capture small multinomial probabilities with sufficient regularity. These observations are not surprising as Method 1 employs the variance of a multinomial random variable while Method 2 utilizes an estimator of the variance (equations 2 and 4).

The simulation results serve as an example and a numerical verification of the expected performance of the two methods. As might be expected, Method 1 produces simultaneous confidence intervals with better statistical properties than does

Method 2. In some cases the differences are not substantial, however they are extremely consistent. In light of these results, it is surprising that Method 1 is not more widely employed. Although Method 1 is slightly more complex than Method 2, the formulas are not a substantial burden and can be easily performed on a hand-held calculator. In addition, confidence intervals for a number of discrete distributions can be obtained through the approach embodied in equation (2) and would likely produce confidence intervals with properties equal or superior to more frequently used methods.

REFERENCES

Borland International. (1988). *Turbo C User's Guide*. Scott's Valley, CA: Borland International.

Goodman, L.A. (1965). On simultaneous confidence intervals for multinomial proportions. *Technometrics* 7, 247-254.

Johnson, R. and Bhattacharyya, G. (1987). *Statistics: Principles and Methods*. New York: John Wiley and Sons.

Kernighan, B. and Ritchie, D. (1988). *The C Programming Language*, 1st edition. Englewood-Cliffs, NJ: Prentice-Hall.

Press, W.H., Flannery, B.P., Teukolsky, S.A., and Vetterling, W.T. (1986). *Numerical Recipes: The Art of Scientific Computing*. New York: Cambridge University Press.

Kennedy, W. and Gentle, J. (1980). *Statistical Computing*. New York: Marcel Dekker.

Seber, G.A.F. (1982). *The Estimation of Animal Abundance*, 2nd edition. New York: Macmillan.

Sokal, R.R., and Rohlf, F.J. (1987). *Introduction to Biostatistics*, 2nd edition. New York: W.H. Freeman and Company.

Table 1. The multinomial distributions from which samples were drawn.

Distribution	# Categories	Probabilities
1	2	equal
2	2	0.2, 0.8
3	3	equal
4	3	0.05, 0.30, 0.65
5	5	equal
6	5	0.05, 0.10, 0.20, 0.30, 0.35
7	7	equal
8	10	equal

Table 2. Simultaneous coverage observed in 5000 independent 95% simultaneous confidence intervals.

Distribution	n	Method 1	Method 2	n	Method 1	Method 2
1	25	0.9568	0.9568	50	0.9284	0.9284
1	75	0.9346	0.9346	100	0.9436	0.9436
1	150	0.9592	0.9394	200	0.9476	0.9476
1	300	0.9302	0.9302	500	0.9460	0.9460
2	25	0.9319	0.8940	50	0.9536	0.9448
2	75	0.9444	0.9356	100	0.9448	0.9340
2	150	0.9468	0.9402	200	0.9606	0.9422
2	300	0.9510	0.9472	500	0.9464	0.9454
3	25	0.9454	0.9472	50	0.9620	0.9562
3	75	0.9500	0.9420	100	0.9486	0.9424
3	150	0.9466	0.9494	200	0.9464	0.9462
3	300	0.9562	0.9542	500	0.9406	0.9384
4	25	0.6691	0.6849	50	0.8930	0.8972
4	75	0.9502	0.8704	100	0.9616	0.9380
4	150	0.9572	0.9142	200	0.9634	0.9422
4	300	0.9616	0.9392	500	0.9582	0.9462
5	25	0.9542	0.8664	50	0.9656	0.9124
5	75	0.9296	0.8864	100	0.9556	0.9330
5	150	0.9524	0.9470	200	0.9530	0.9546
5	300	0.9556	0.9470	500	0.9706	0.9616
6	25	0.6508	0.6184	50	0.8820	0.8416
6	75	0.9216	0.8090	100	0.9528	0.9142
6	150	0.9540	0.8950	200	0.9460	0.9262
6	300	0.9620	0.9486	500	0.9606	0.9372
7	25	0.8342	0.8626	50	0.9360	0.8720
7	75	0.9380	0.9230	100	0.9480	0.8764
7	150	0.9430	0.9218	200	0.9438	0.9520
7	300	0.9728	0.9504	500	0.9804	0.9794
8	25	0.4140	0.4344	50	0.8648	0.6848
8	75	0.9408	0.8464	100	0.9608	0.7644
8	150	0.9682	0.8748	200	0.9484	0.9228
8	300	0.9726	0.9146	500	0.9668	0.9474

Table 3. Individual coverage and average confidence interval width observed in 5000 independent 95% simultaneous confidence intervals.

Distribution	Probability	n	Coverage Method 1	Coverage Method 2	Width Method 1	Width Method 2
1	0.50	25	0.9568	0.9568	0.3585	0.3919
1	0.50	50	0.9284	0.9284	0.2645	0.2771
1	0.50	75	0.9346	0.9346	0.2193	0.2263
1	0.50	100	0.9436	0.9436	0.1914	0.1960
1	0.50	150	0.9592	0.9394	0.1575	0.1601
1	0.50	200	0.9476	0.9476	0.1369	0.1386
1	0.50	300	0.9302	0.9302	0.1123	0.1132
1	0.50	500	0.9460	0.9460	0.0872	0.0877
2	0.20	25	0.9319	0.8940	0.2957	0.3058
2	0.20	50	0.9536	0.9448	0.2151	0.2205
2	0.20	75	0.9444	0.9356	0.1773	0.1804
2	0.20	100	0.9448	0.9340	0.1543	0.1564
2	0.20	150	0.9468	0.9402	0.1266	0.1278
2	0.20	200	0.9606	0.9422	0.1100	0.1107
2	0.20	300	0.9510	0.9472	0.0900	0.0904
2	0.20	500	0.9464	0.9454	0.6999	0.0701
3	1/3	25	0.9805	0.9800	0.4045	0.4487
3	1/3	50	0.9859	0.9836	0.3013	0.3188
3	1/3	75	0.9811	0.9781	0.2506	0.2604
3	1/3	100	0.9811	0.9785	0.2191	0.2255
3	1/3	150	0.9803	0.9809	0.1806	0.1842
3	1/3	200	0.9806	0.9795	0.1572	0.1595
3	1/3	300	0.9827	0.9821	0.1290	0.1303
3	1/3	500	0.9763	0.9754	0.1003	0.1009
4	0.05	25	0.6855	0.7197	0.1934	0.1362
4	0.05	50	0.9138	0.9230	0.1559	0.1182
4	0.05	75	0.9660	0.8934	0.1283	0.1060
4	0.05	100	0.9846	0.9624	0.1104	0.0970
4	0.05	150	0.9788	0.9364	0.0887	0.0828
4	0.05	200	0.9876	0.9682	0.0761	0.0728
4	0.05	300	0.9898	0.9692	0.0616	0.0598
4	0.05	500	0.9878	0.9778	0.0473	0.0465
4	0.30	25	0.9718	0.9578	0.3890	0.4261
4	0.30	50	0.9808	0.9770	0.2923	0.3081
4	0.30	75	0.9896	0.9786	0.2438	0.2527
4	0.30	100	0.9820	0.9800	0.2133	0.2192
4	0.30	150	0.9850	0.9832	0.1758	0.1790
4	0.30	200	0.9806	0.9828	0.1529	0.1551
4	0.30	300	0.9820	0.9784	0.1255	0.1266
4	0.30	500	0.9784	0.9776	0.0976	0.0981

Table 3 (continued). Individual coverage and average confidence interval width observed in 5000 independent 95% simultaneous confidence intervals.

Distribution	Probability	n	Coverage Method 1	Coverage Method 2	Width Method 1	Width Method 2
4	0.65	25	0.9782	0.9546	0.4019	0.4464
4	0.65	50	0.9830	0.9786	0.3031	0.3211
4	0.65	75	0.9886	0.9816	0.2530	0.2632
4	0.65	100	0.9860	0.9854	0.2215	0.2282
4	0.65	150	0.9838	0.9836	0.1827	0.1864
4	0.65	200	0.9820	0.9786	0.1590	0.1614
4	0.65	300	0.9814	0.9804	0.1305	0.1318
4	0.65	500	0.9764	0.9752	0.1015	0.1021
5	0.20	25	0.9900	0.9722	0.3787	0.3852
5	0.20	50	0.9930	0.9818	0.2793	0.2891
5	0.20	75	0.9852	0.9761	0.2312	0.2370
5	0.20	100	0.9910	0.9861	0.2016	0.2054
5	0.20	150	0.9902	0.9889	0.1658	0.1679
5	0.20	200	0.9903	0.9905	0.1441	0.1455
5	0.20	300	0.9909	0.9888	0.1181	0.1189
5	0.20	500	0.9940	0.9921	0.0917	0.0921
6	0.05	25	0.7184	0.7286	0.2117	0.1442
6	0.05	50	0.9136	0.9242	0.1699	0.1236
6	0.05	75	0.9604	0.8918	0.1392	0.1112
6	0.05	100	0.9824	0.9632	0.1199	0.1024
6	0.05	150	0.9904	0.9398	0.0962	0.0884
6	0.05	200	0.9886	0.9692	0.0824	0.0781
6	0.05	300	0.9914	0.9842	0.0665	0.0643
6	0.05	500	0.9948	0.9816	0.0510	0.0500
6	0.10	25	0.9218	0.9286	0.2968	0.2427
6	0.10	50	0.9872	0.9666	0.2209	0.2012
6	0.10	75	0.9916	0.9514	0.1805	0.1732
6	0.10	100	0.9896	0.9768	0.1560	0.1525
6	0.10	150	0.9894	0.9854	0.1271	0.1254
6	0.10	200	0.9890	0.9856	0.1099	0.1088
6	0.10	300	0.9946	0.9878	0.0896	0.0890
6	0.10	500	0.9940	0.9890	0.0693	0.0690
6	0.20	25	0.9838	0.9690	0.3748	0.3817
6	0.20	50	0.9928	0.9808	0.2786	0.2884
6	0.20	75	0.9878	0.9778	0.2309	0.2368
6	0.20	100	0.9928	0.9904	0.2016	0.2055
6	0.20	150	0.9890	0.9874	0.1658	0.1679
6	0.20	200	0.9852	0.9884	0.1441	0.1455
6	0.20	300	0.9928	0.9932	0.1181	0.1189
6	0.20	500	0.9946	0.9934	0.0918	0.0921

Table 3 (continued). Individual coverage and average confidence interval width observed in 5000 independent 95% simultaneous confidence intervals.

Distribution	Probability	n	Coverage Method 1	Coverage Method 2	Width Method 1	Width Method 2
6	0.30	25	0.9918	0.9590	0.4162	0.4595
6	0.30	50	0.9894	0.9798	0.3131	0.3322
6	0.30	75	0.9888	0.9868	0.2612	0.2720
6	0.30	100	0.9910	0.9894	0.2287	0.2358
6	0.30	150	0.9906	0.9864	0.1887	0.1926
6	0.30	200	0.9874	0.9876	0.1643	0.1668
6	0.30	300	0.9914	0.9920	0.1349	0.1363
6	0.30	500	0.9898	0.9874	0.1049	0.1056
6	0.35	25	0.9892	0.9806	0.4303	0.4837
6	0.35	50	0.9902	0.9764	0.3245	0.3462
6	0.35	75	0.9876	0.9890	0.2711	0.2833
6	0.35	100	0.9950	0.9906	0.2374	0.2456
6	0.35	150	0.9930	0.9916	0.1961	0.2006
6	0.35	200	0.9926	0.9922	0.1707	0.1737
6	0.35	300	0.9886	0.9882	0.1402	0.1419
6	0.35	500	0.9852	0.9832	0.1091	0.1099
7	1/7	25	0.9741	0.9796	0.3551	0.3216
7	1/7	50	0.9907	0.9810	0.2609	0.2566
7	1/7	75	0.9909	0.9889	0.2146	0.2152
7	1/7	100	0.9925	0.9820	0.1865	0.1872
7	1/7	150	0.9917	0.9886	0.1528	0.1532
7	1/7	200	0.9919	0.9931	0.1325	0.1328
7	1/7	300	0.9961	0.9928	0.1084	0.1085
7	1/7	500	0.9972	0.9970	0.0841	0.0841
8	0.10	25	0.9183	0.9277	0.3234	0.2564
8	0.10	50	0.9854	0.9652	0.2418	0.2134
8	0.10	75	0.9940	0.9840	0.1975	0.1863
8	0.10	100	0.9960	0.9750	0.1705	0.1653
8	0.10	150	0.9967	0.9873	0.1388	0.1366
8	0.10	200	0.9948	0.9920	0.1200	0.1185
8	0.10	300	0.9973	0.9914	0.0977	0.0970
8	0.10	500	0.9967	0.9947	0.0756	0.0752

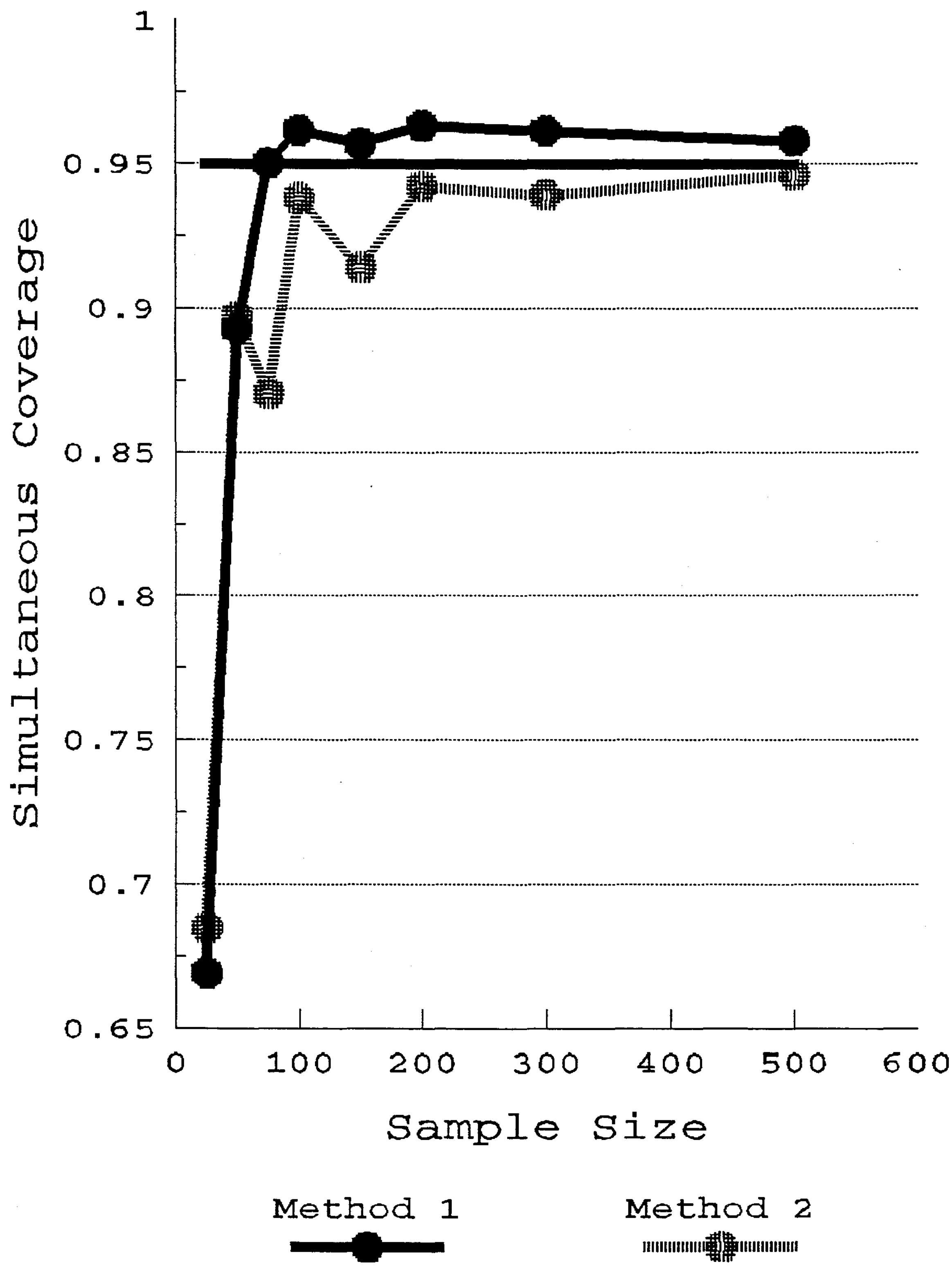

Figure 1. Simultaneous coverage observed in 95% simultaneous confidence intervals constructed about the probabilities of distribution 4.

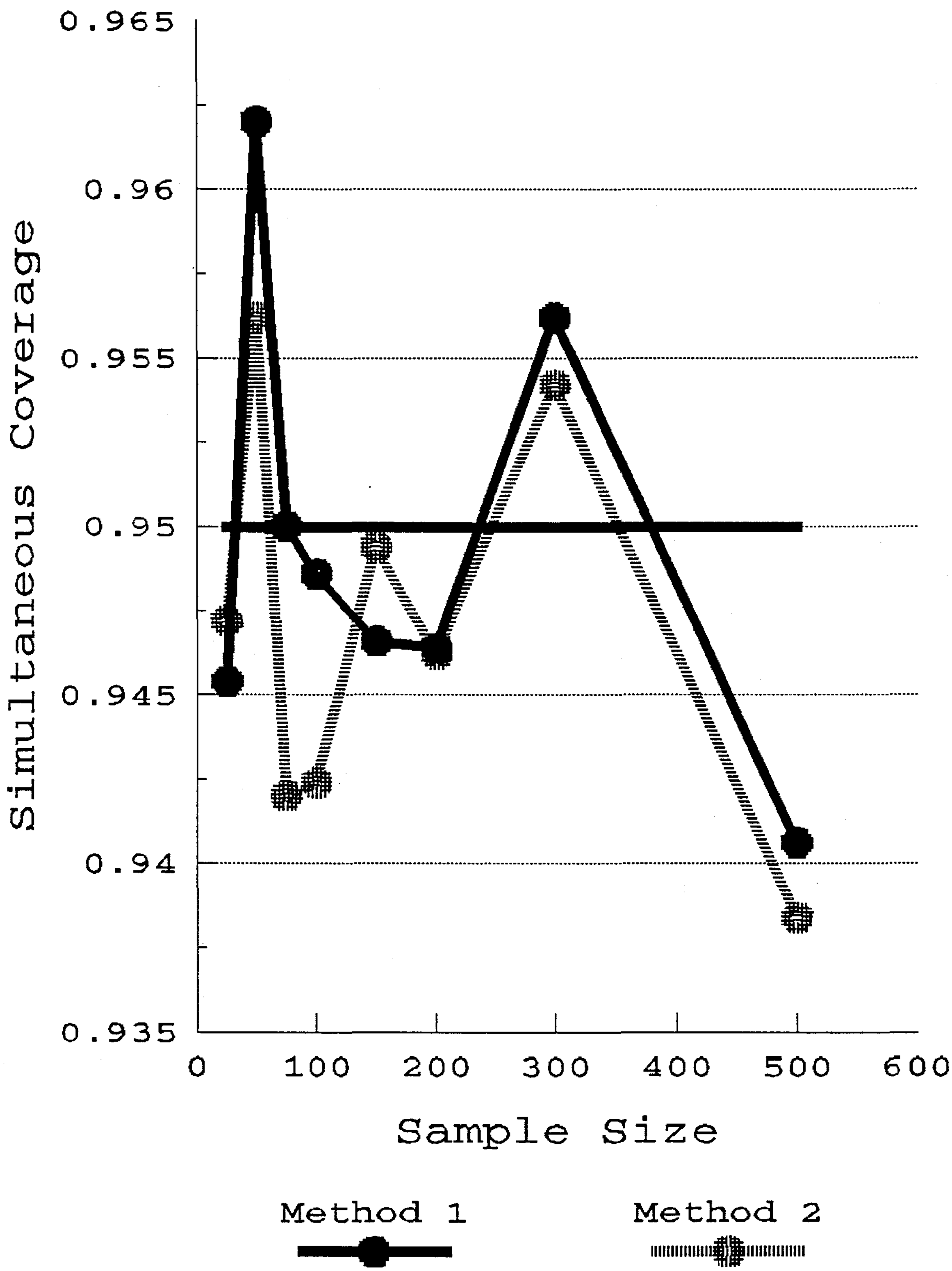

Figure 2. Simultaneous coverage observed in 95% simultaneous confidence intervals constructed about the probabilities of distribution 3.

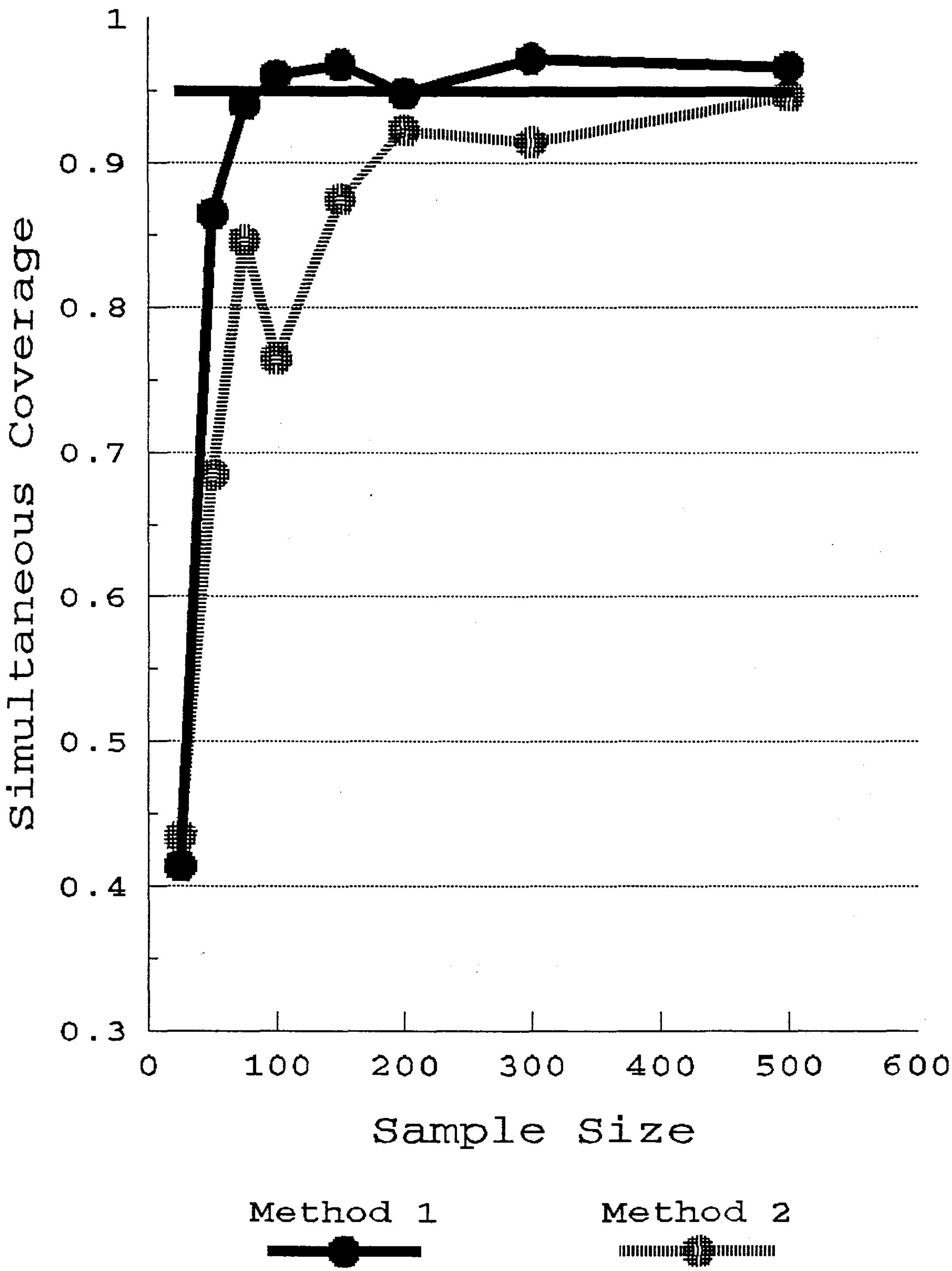

Figure 3. Simultaneous coverage observed in 95% simultaneous confidence intervals constructed about the probabilities of distribution 8.

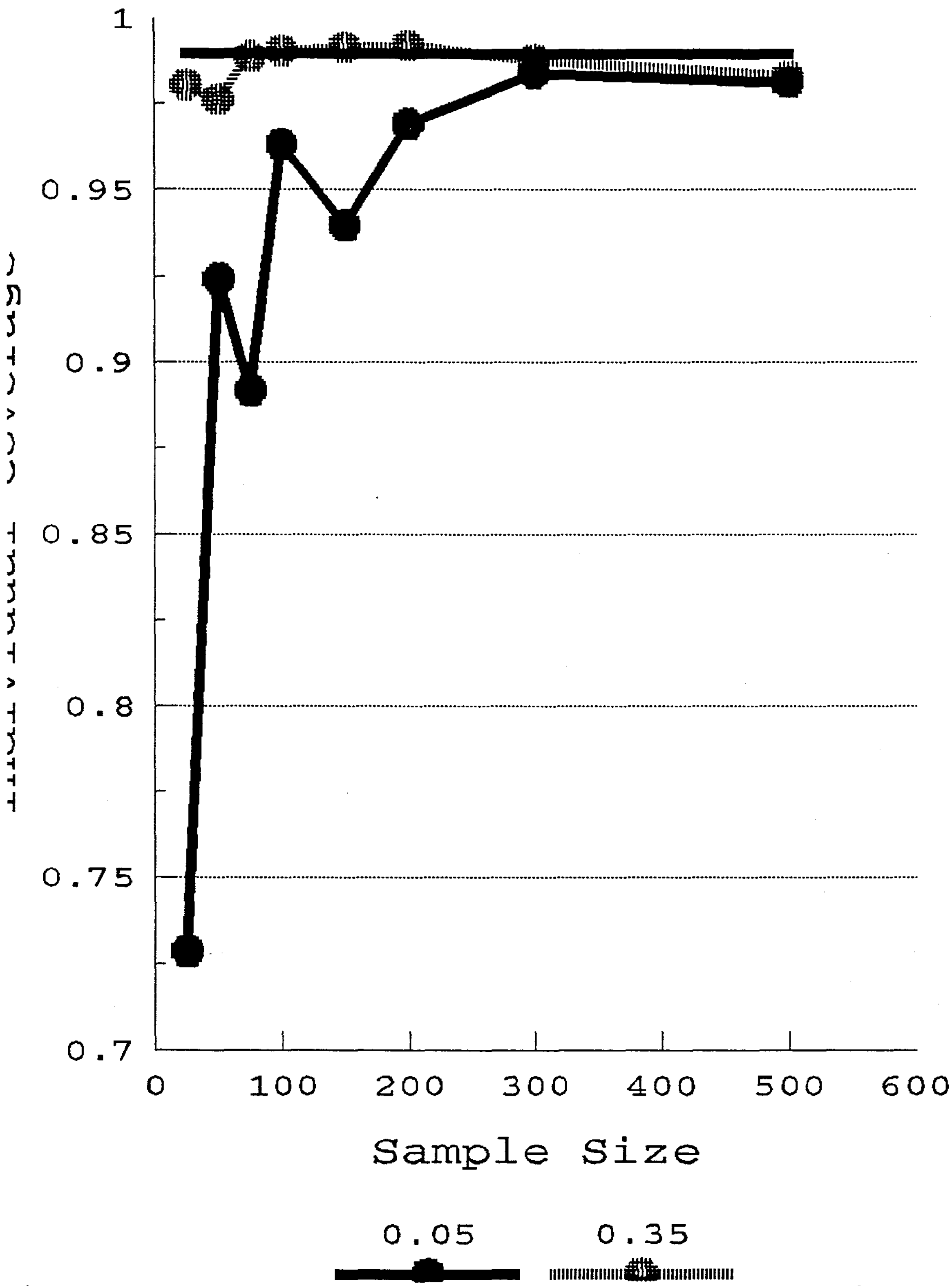

Figure 4. Individual coverage observed in 99% confidence intervals constructed about the probabilities 0.05 and 0.35 of distribution 6 using Method 2.

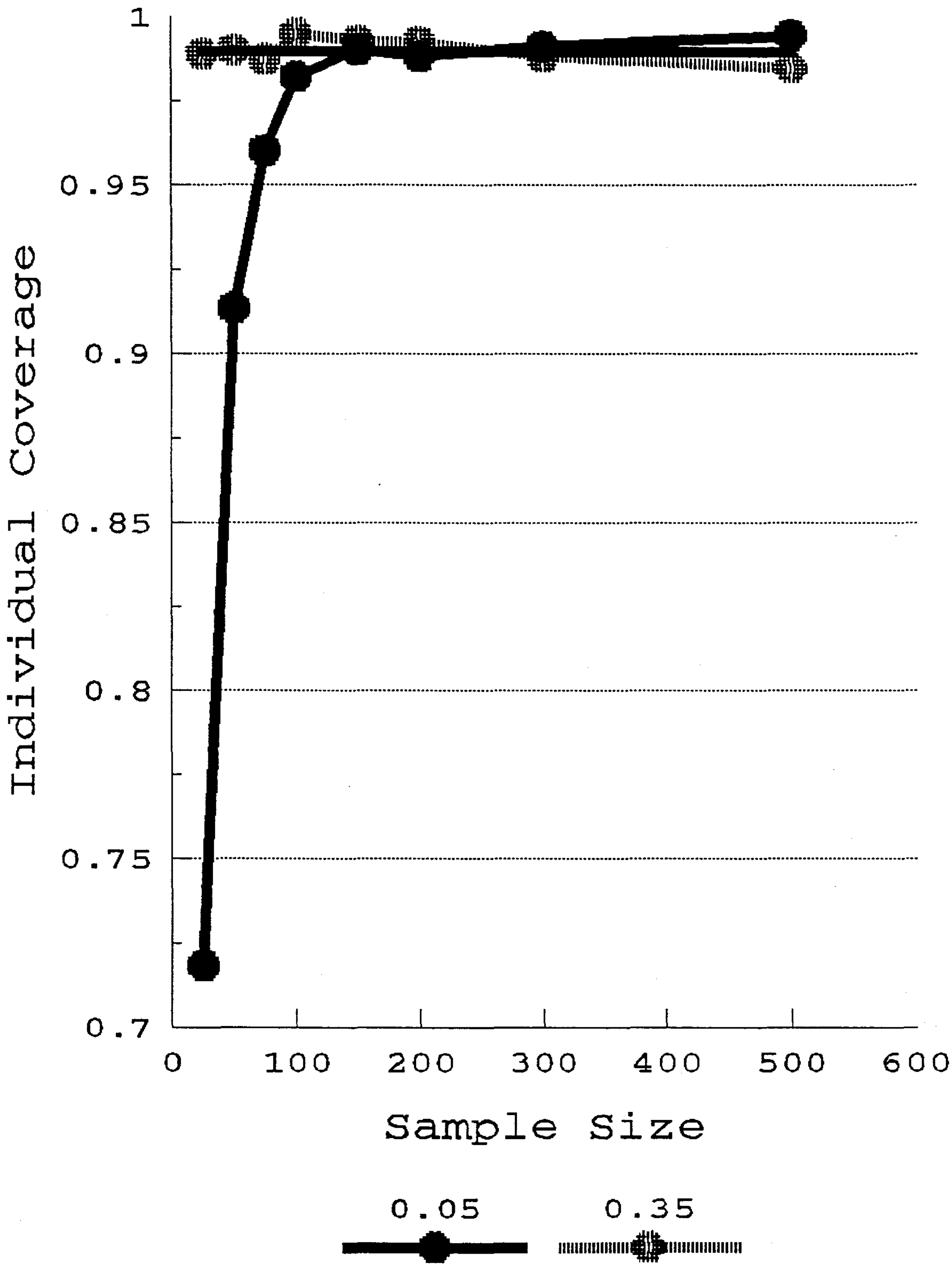

Figure 5. Individual coverage observed in 99% confidence intervals constructed about the probabilities 0.05 and 0.35 of distribution 6 using Method 1.

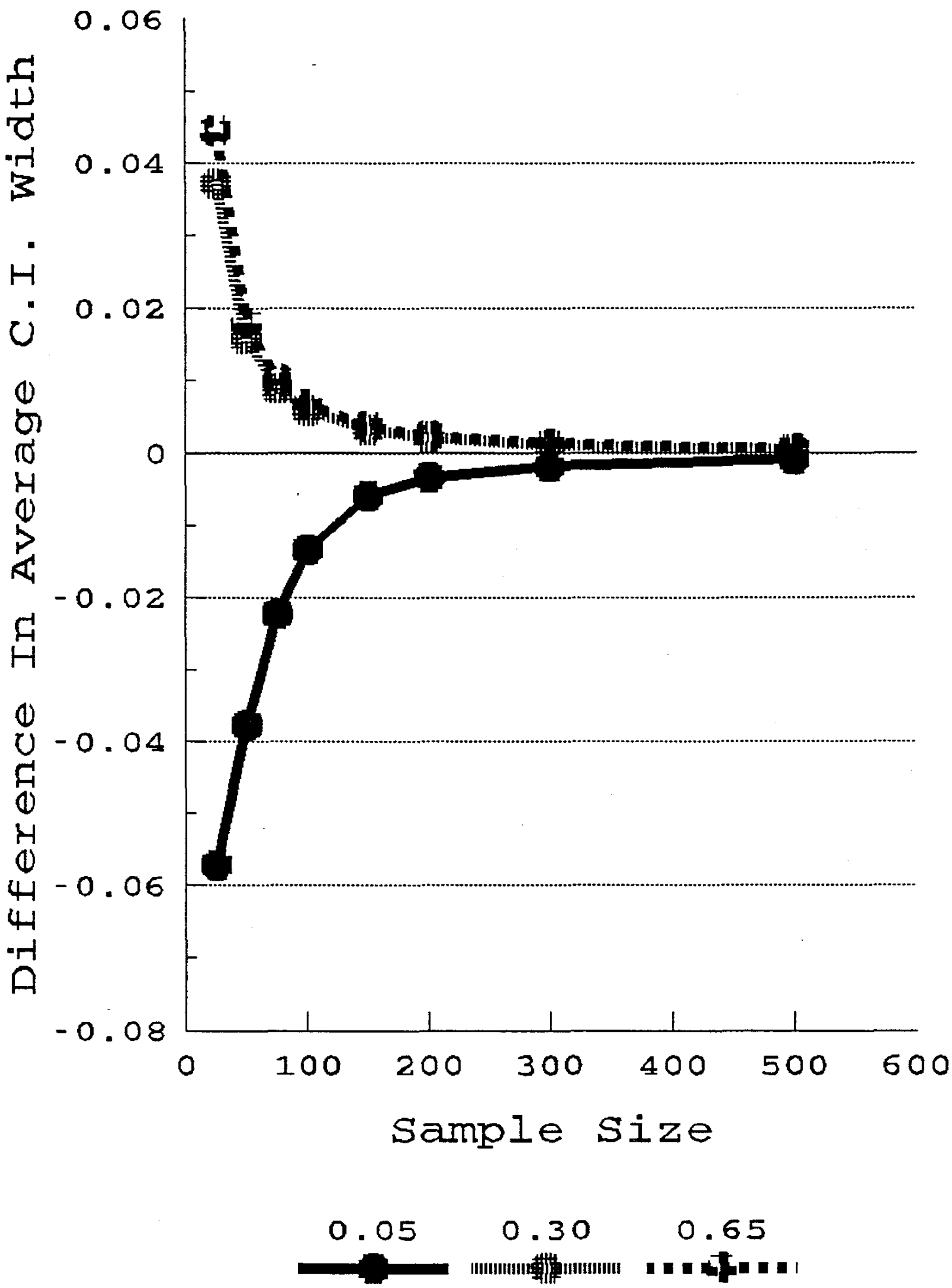

Figure 6. The difference in average confidence interval widths (Method 2 - Method 1) observed in 95% simultaneous confidence intervals constructed about the probabilities of distribution 4.